NATIONAL QUALITY FORUM
Driving measurable health improvements together

Meeting Summary

All-Cause Admissions and Readmissions Standing Committee – Measure Evaluation Web Meeting

The National Quality Forum (NQF) convened the All-Cause Admissions and Readmissions Standing Committee for a web meeting on July 6, 2021, to evaluate four measures.

Welcome, Introductions, and Review of Meeting Objectives

NQF welcomed the Standing Committee and participants to the web meeting. NQF staff reviewed the meeting objectives. The Standing Committee members each introduced themselves and disclosed any conflicts of interest. No conflicts were disclosed.

Some Standing Committee members were unable to attend the entire meeting due to early departures and late arrivals. The vote totals reflect members present and eligible to vote. Quorum was met and maintained for the entirety of the meeting.

Topic Area Introduction and Overview of Evaluation Process

NQF staff provided an overview of the topic area and the current NQF portfolio of endorsed measures. Additionally, NQF reviewed the Consensus Development Process (CDP) and the measure evaluation criteria.

Measure Evaluation

During the meeting, the All-Cause Admissions and Readmissions Standing Committee evaluated four measures, including three maintenance measures and one new measure for endorsement consideration. A summary of the Standing Committee's deliberations will also be compiled and provided in the draft technical report. NQF will post the draft technical report on August 19, 2021, for public comment on the NQF website. The draft technical report will be posted for 30 calendar days.

Rating Scale: H – High; M – Medium; L – Low; I – Insufficient; NA – Not Applicable

#2860 30-Day All-Cause Unplanned Readmission Following Psychiatric Hospitalization in an Inpatient Psychiatric Facility (IPF) – (Mathematica/Centers for Medicare & Medicaid Services [CMS])

Measure Steward/Developer Representatives at the Meeting

Jason Smoot

Standing Committee Votes

- **Evidence**: Pass-15; No Pass-0
- **Performance Gap**: H-1; M-11; L-2; I-0
- **Reliability**: Yes-14; No-0
    - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.
    - The NQF Scientific Methods Panel's rating for Reliability: Moderate (H-0; M-8; L-1; I-0)

    - The Standing Committee accepted the NQF Scientific Methods Panel's rating.
- **Validity**: Yes-14; No-0
    - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.
    - The NQF Scientific Methods Panel's rating for Validity: Moderate (H-1; M-6; L-1; I-1)
    - The Standing Committee accepted the NQF Scientific Methods Panel's rating.
- **Feasibility**: H-6; M-8; L-0; I-0
- **Use**: Pass-14; No Pass-0
- **Usability**: H-1; M-14; L-0; I-0

Standing Committee Recommendation for Endorsement: Yes-15; No-0 (Pass – 15/15)

The Standing Committee recommended this outcome measure for continued endorsement.

The developer provided a brief description of the measure, with specific reference to the principal diagnoses under examination and the 24-month performance period. The Standing Committee discussed the developer's inclusion of new evidence for this measure, which demonstrated an association of various hospital- and facility-led interventions that can be implemented to improve this outcome. A Standing Committee member noted the importance of considering other evidence that may be more patient-centric, such as lifestyle changes and nonmedicinal interventions. The same Standing Committee member expressed recognition of the vital role of medications in psychiatric medicine, but they further stated that there are good reasons to study alternative therapies. The Standing Committee did not have any additional commentary or concerns with regard to the evidence, and it passed the measure unanimously on this criterion.

The Standing Committee reflected on the interpretation of the performance gap, noting that although the margins of improvement are low and the decreased unplanned readmission rate is not statistically significant, change is evident, improvement is observable, and continuation of the measure is important. A Standing Committee member expressed that even in the smallest margins of improvement, improvement is quite significant in relation to psychiatric illness due to the challenges of managing the condition. Another Standing Committee member expressed that the current measure information, as it is presented, does not necessarily allow for full transparency to determine whether all providers are improving and would like to see that transparency in future measure submissions. The Standing Committee acknowledged that the data clearly show the existence of disparities. The Standing Committee expressed no additional concerns and passed the measure on the performance gap criterion.

Moving to scientific acceptability (i.e., reliability and validity), the Standing Committee reviewed the testing information. The developer used two techniques to demonstrate measure score reliability, namely split sampling and bootstrapping. The Standing Committee reiterated the split sample intra-class correlation coefficient (ICC) value of 0.559 and the bootstrapping method's ICC value of 0.752. The Standing Committee acknowledged that the NQF Scientific Methods Panel (SMP) reviewed and passed the measure on reliability. The Standing Committee did not express concern and agreed to uphold the SMP's rating of "moderate" on the reliability criterion.

With respect to validity, the Standing Committee noted both approaches that the developer used to conduct validity testing and reiterated the testing results of both techniques that were used. Those approaches were the Spearman rank correlation and discriminant validity testing. Discriminant validity was tested against six patient characteristics hypothesized to be associated with higher readmissions rates: (1) male patients, (2) patients with a substance use disorder (SUD), (3) patients with schizophrenia, (4) non-White patients, (5) patients with a shorter length of stay at the Inpatient

Psychiatric Facility (IPF), and (6) patients with socioeconomic characteristics associated with worse health outcomes. The results ranged from 0.012 to 0.457 and 0.05 to 0.473 for predicted and expected rates, respectively. For the observed rates, the results were smaller, ranging from 0.003 to 0.109. The Standing Committee noted the small range of the discriminant validity testing results; nevertheless, it recognized that validity is still demonstrated and agreed to uphold the SMP's rating of moderate on the validity criterion.

The Standing Committee noted the claims-based nature of the measure and did not express any concern with respect to the feasibility of the measure. Therefore, the Standing Committee voted to pass the measure on feasibility. Moving to use and usability, the Standing Committee recognized that this measure is used in the Centers for Medicare & Medicaid Services' (CMS) Inpatient Psychiatric Facility Prospective Payment System (IPF PPS) program. The Standing Committee also referenced the coinciding national IPF readmission rate, noting that it decreased from 20.1 percent to 18.5 percent between 2017 and 2019. One Standing Committee member offered a patient-centric observation of the usefulness of the hospital rating system, noting that patients are appreciative and find it useful. The Standing Committee did not raise any major concerns and passed the measure on the use and usability criteria.

Prior to the overall suitability for endorsement vote, a Standing Committee member shared some concerns about the 30-day reporting period. The Standing Committee member called into question the effectiveness of measures bound by a 30-day transition period and stated that in relation to unexpected or unintended harms, the data may mislead people to a false sense of assuredness that the care coordination is better than it may be after 30 days. An additional Standing Committee member commented on the 30-day period, describing it as arbitrary, and recommended that the 30 days be paired with an indication of adequate follow-up. Another Standing Committee member recommended that CMS and the developers consider assessing longer time periods in order to determine whether there is a statistically significant difference in performance over a longer duration of observation time. The Standing Committee acknowledged that an extension to the 30-day window cannot necessarily be implemented immediately, and such changes would likely convert the measure into a new one. The developer responded by stating that the premise behind the 30-day period was to be consistent with readmissions measures that have been endorsed and are publicly reported in CMS' inpatient public reporting programs.

The Standing Committee then voted to recommend the measure for continued endorsement. The Standing Committee observed that there are several related measures to this metric, but it did not consider these measures to be competing.

#2880 Excess Days in Acute Care (EDAC) After Hospitalization for Heart Failure (HF) – (Yale Center for Outcomes Research & Evaluation [CORE])

Measure Steward/Developer Representatives at the Meeting

Jackie Grady

Doris Peter

Lisa Suter

Jim Poyer (CMS)

Standing Committee Votes

- **Evidence**: Pass-14; No Pass-1
- **Performance Gap**: H-4; M-7; L-4; I-0

- **Reliability:** Yes-15; No-0
    - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.
    - The NQF Scientific Methods Panel's rating for Reliability: Moderate (H-0; M-8; L-1; I-0)
    - The Standing Committee accepted the NQF Scientific Methods Panel's rating.
- **Validity**: Yes-11; No-4
    - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.
    - The NQF Scientific Methods Panel's rating for Validity: Moderate (H-0; M-7; L-0; I-1)
    - The Standing Committee accepted the NQF Scientific Methods Panel's rating.
- **Feasibility**: H-5; M-10; L-0; I-0
- **Use**: Pass-15; No Pass-0
- **Usability**: H-2; M-9; L-3; I-1

Standing Committee Recommendation for Endorsement: Yes-15; No-0 (Pass – 15/15)

The Standing Committee recommended this outcome measure for continued endorsement.

The developer presented an overview of the measure, noting that this and other Excess Days in Acute Care (EDAC) measures were intended to be an expansion of other mortality and readmission measures for conditions such as heart failure (HF) and pneumonia. Measures of unplanned readmissions already exist; nonetheless, these EDAC measures target emergency department (ED) and observation stay utilization for these conditions. The developer posits that by capturing a range of acute care events that are important to patients, a more complete picture of post-discharge outcomes can be identified. This EDAC measure is intended to be a measure of the days spent in acute care within 30 days of discharge from an inpatient hospitalization for HF. Specifically, it is an outcome measure that is intended to capture the quality of care transitions provided to discharged patients who had an HF hospitalization by collectively measuring a set of adverse acute care outcomes that can occur post-discharge: ED visits, observation stays, and unplanned readmissions at any time during the 30 days post-discharge.

The Standing Committee reviewed and discussed the evidence, noting that the developer cited several studies supporting various care processes that can influence post-discharge acute care utilization after a hospitalization for HF. Further, the developer provided evidence suggesting that hospitals and health plans have been able to reduce readmission rates through more generalizable quality improvement initiatives, such as communication between providers, patient education, patient safety, and coordinated transitions to the outpatient environment. One Standing Committee member asked whether the current readmission measures for conditions such as HF and pneumonia would be retired, as the EDAC measures have a more wholistic capture of utilization. Jim Poyer, CMS representative, replied that the readmission measures are required by statute within the Hospital Readmissions Reduction Program (HRRP); therefore, the current intent is to use both EDAC and readmission measures in parallel until a change occurs in the statute. In addition, the Standing Committee asked whether the developer has any evidence to show that the number of days, rather than readmissions, can be influenced by the behavior in the initial acute hospitalization. The developer responded by stating that they have anecdotal evidence showing that this measure has an impact on the number of days. The Standing Committee did not raise any other questions or concerns and passed the measure on the evidence criterion.

Moving to the performance gap criterion, the Standing Committee considered the range of performance across hospitals with at least 25 admissions during the reporting period of 2016–2019, in which the rates ranged from -59.7 to 154.4 EDAC per 100 admissions with a median EDAC of 2.3 per 100 admissions. The Standing Committee also acknowledged that this measure was able to identify disparities, namely for

dual-eligible patients and by using the Agency for Healthcare Research and Quality (AHRQ) Socioeconomic Status (SES) Index. One Standing Committee member asked whether a comparison was done to see whether the variation in performance with the current EDAC measure was more or less than that of the HF readmissions measure. The developer stated that the EDAC measures do have a wider range of days; however, this is not entirely analogous or comparable to the readmission measures due to how the measure is calculated. The Standing Committee had no further questions and voted to pass the measure on performance gap.

The Standing Committee then reviewed the scientific acceptability of the measure. The Standing Committee acknowledged that the SMP reviewed and passed this measure on reliability with a moderate rating. The Standing Committee also noted that the developer conducted testing at the measure score level and calculated an ICC using a split-sample approach. The developer reported ICC ranges from 0.456 for hospitals with at least two admissions to 0.698 for hospitals with at least 300 admissions. For hospitals with at least 25 admissions, the ICC was 0.527. The Standing Committee did not raise any questions or concerns and upheld the SMP's rating for reliability.

The Standing Committee also acknowledged that the SMP passed the measure on validity with a moderate rating. The Standing Committee considered the validity testing for this measure, noting that the developer conducted face validity testing via a Technical Expert Panel (TEP) as well as empirical validity testing. The developer reported that 11 out of 12 (91.7 percent) TEP members convened by the developer strongly, moderately, or somewhat agreed with the following statement: "The risk-standardized acute care days obtained from the measures as specified can be used to distinguish between better and worse quality hospitals." The developer also conducted construct validity testing to determine the relationships between the HF EDAC measure score and the risk-standardized readmissions rate group scores, the overall hospital rating scores, and the HF readmissions measure. The developer reported statistically significant correlations for all measures in the direction hypothesized. For the risk adjustment model, the Standing Committee noted that two social risk factors were tested and found to be statistically significant (i.e., dual-eligible status and AHRQ SES index). The developer also performed a decomposition analysis. In this analysis, the clinical risk factors have a larger patient-level effect compared with their hospital-level effects. In contrast, both the low AHRQ SES variable and the dual-eligible variable have a larger hospital-level effect compared with the patient-level effect. Based on these analyses, the developer did not adjust this measure for either dual eligibility or the AHRQ SES Index. One Standing Committee member raised concern with the low r-squared value of 0.027. The developer replied that when looking at count models, such as the number of days, the goal is to adjust for the case-mix rather than to predict an outcome. As a result, the deviance r-squared value does not have the same interpretation as a prediction, and the results are what would be expected with this type of model and for this type of data.

The Standing Committee asked for more explanation regarding their rationale for excluding the social risk factors in the final model, namely the decomposition analysis. The developer replied that the decomposition analysis evaluates the variation that can be attributed to the hospital and what variation can be attributed to the patient. In this analysis, the clinical risk factors have a larger patient-level effect compared with their hospital-level effects. In contrast, both the low AHRQ SES variable and the dual-eligible variable have larger hospital-level effects compared with the patient-level effect. Based on these analyses, the developer did not adjust this measure for either dual eligibility or the AHRQ SES Index. Mr. Poyer also commented that CMS does not adjust for social risk factors, such as dual eligibility, at the measure level. Rather, the HRRP, in which most of the measures are currently used, stratifies its payment calculations in accordance with statutory guidance based on dual eligibility. One Standing Committee commented that having the decomposition analysis was helpful in better understanding the

exclusion of the social risk factors. Another Standing Committee member commented that these factors should be included, regardless of the magnitude of their effect, such that there are unbiased assignments and reporting of accountability. Co-Chair John Bulger asked NQF whether there is any current work being done at NQF to address the concerns regarding social risk factor adjustment within quality measurement. In response, Dr. Matt Pickering, NQF senior director, stated that NQF is currently developing technical guidance, which has been released for public comment. This guidance is intended to provide a step-by-step approach for social and functional status-related risk adjustment within quality measurement. This guidance will help to evolve NQF's current criteria, which will occur after 2022. Therefore, measures under review for the spring 2021 cycle must be evaluated under NQF's current criteria. The Standing Committee raised no other questions or concerns and voted to uphold the SMP's rating for validity.

Moving to feasibility, the Standing Committee acknowledged that this measure uses administrative claims and enrollment data, and as such, it offers no data collection burden to hospitals or providers. The Standing Committee did not raise any questions or concerns and passed the measure on the feasibility criterion. The Standing Committee then evaluated the use criterion and noted that the measure is currently used within the Hospital Inpatient Quality Reporting Program (HIQRP) and Care Compare. The Standing Committee also noted that the developer routinely scans the literature for articles describing research related to this measure and solicits feedback from stakeholders. The Standing Committee did not have any questions or concerns and passed the measure on the use criterion.

For usability, the Standing Committee recognized that the developer reported improvement over the past three reporting periods (i.e., 2014–2017, 2015–2018, and 2016–2019) in measure scores across most of the distribution, from the 30th through the 80th percentile. One Standing Committee member questioned whether this measure is contributing valuable information for hospitals or whether it is simply adding noise to the system, as there are various readmission measures currently used within Hospital Compare. This Standing Committee member also questioned which of the measures hospitals should target, namely whether hospitals should target the number of excess days or whether a readmission occurred. The developer replied that when the readmission measures were first developed, various stakeholders and members of the public expressed a great deal of concern that hospitals may game the system with these measures (i.e., increased observation stays). Therefore, CMS, in response to these concerns, developed the EDAC measures. CMS has implemented the EDAC measures as balancing measures, and hospitals that continue to use these measures have expressed value in these measures. Another Standing Committee member added that they do see the value in this measure but would like to see an excess days measure that is cross-cutting and not simply condition-specific. One Standing Committee member agreed that this measure has increased usability due to the value of monitoring excess days in acute care, such as looking at observations stays. The Standing Committee also recognized that the developer continues to monitor for unintended consequences. There were no other questions or concerns raised, and the Standing Committee passed the measure on usability.

The Standing Committee then voted to recommend the measure for continued endorsement. The Standing Committee observed that there are several related measures to this metric, but it did not consider these measures to be competing.

#2882 Excess Days in Acute Care (EDAC) After Hospitalization for Pneumonia - (Yale CORE)

Measure Steward/Developer Representatives at the Meeting

Jackie Grady

Doris Peter

Lisa Suter

Jim Poyer (CMS)

Standing Committee Votes

- **Evidence**: Pass-15; No Pass-1
- **Performance Gap:** H-1; M-14; L-0; I-0
- **Reliability**: Yes-14; No-1
    - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.
    - The NQF Scientific Methods Panel's rating for Reliability: Moderate (H-1; M-8; L-0; I-0)
    - The Standing Committee accepted the NQF Scientific Methods Panel's rating.
- **Validity**: Yes-14; No-1
    - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.
    - The NQF Scientific Methods Panel's rating for Validity: Moderate (H-0; M-7; L-0; I-1)
    - The Standing Committee accepted the NQF Scientific Methods Panel's rating.
- **Feasibility**: H-7; M-8; L-0; I-0
- **Use**: Pass-15; No Pass-0
- **Usability**: H-0; M-11; L-2; I-2

Standing Committee Recommendation for Endorsement: Yes-13; No-2 (Pass – 13/15)

The Standing Committee recommended this outcome measure for continued endorsement.

The developer presented an overview of the measure, noting once again that this and other EDAC measures were intended to be an expansion of other mortality and readmission measures for conditions such as HF and pneumonia. This measure assesses the days spent in acute care within 30 days of discharge from an inpatient hospitalization for pneumonia. The developer posits that by capturing a range of acute care events that are important to patients, a more complete picture of post-discharge outcomes can be identified.

The Standing Committee reviewed and discussed the evidence, noting that the developer cited new evidence indicating that pneumonia leads to more than 1 million hospitalizations per year, incurring billions of dollars in healthcare costs. Furthermore, the developer provided a logic model with additional supporting evidence suggesting that hospitals can influence EDAC through a broad range of clinical activities, including communication between providers, patient education, prevention of and response to complications, patient safety, medication reconciliation, better disease management strategies, and coordinated transitions to the outpatient environment. The Standing Committee acknowledged that the same concerns and discussion points raised for NQF #2880 also apply to NQF #2882 and proceeded to pass the measure on the evidence criterion.

The Standing Committee then reviewed performance gap and recognized that the developer reported EDAC scores for the most recent reporting period (2016–2019): -65.7 to 146 EDAC per 100 admissions. The mean was 5.0 EDAC per 100 admissions, and the median was 2.9 EDAC per 100 admissions. The 10th percentile was -23.8, the 50th percentile was 2.9, and the 90th percentile was 36.7 EDAC per 100

admissions. The Standing Committee did not raise any major concerns and passed the measure on the performance gap criterion.

Moving to scientific acceptability, the Standing Committee acknowledged that the SMP reviewed and passed this measure on both reliability and validity. The Standing Committee considered that the developer reported an ICC range of 0.541 for hospitals with at least two admissions to 0.709 for hospitals with at least 300 admissions. For hospitals with at least 25 admissions, the ICC was 0.576. The Standing Committee did not raise any questions and agreed to uphold the SMP's rating for reliability. For validity, the Standing Committee reviewed the validity testing and recognized that both the face validity and empirical validity testing were similar to NQF #2880; thus, the concerns and discussion points raised for NQF #2880 apply to NQF #2882 as well. Therefore, the Standing Committee agreed to accept the SMP's rating for validity.

The Standing Committee did not raise any concerns with respect to feasibility and passed the measure on this criterion. The Standing Committee recognized that the measure is currently used within HIQRP and Care Compare. The Standing Committee also noted that the developer routinely scans the literature for articles describing research related to this measure and solicits feedback from stakeholders. The Standing Committee did not have any questions or concerns and passed the measure on the use criterion. For usability, the Standing Committee noted the minimum improvement in the pneumonia EDAC measure across the three performance periods. The Standing Committee also acknowledged that a contributing factor to the limited improvement could be the severe 2017–2018 influenza season, which would have had an impact on the 2015–2018 and 2016–2019 reporting periods. The Standing Committee also agreed that some of the concerns and discussion points raised for NQF #2880 also apply to NQF #2882. One Standing Committee member mentioned that they would like to see empirical evidence that the use of these EDAC measures carries value compared with the readmission measures, namely whether excess days are meaningful to patients, whether there are correlations to risk-standardized mortality, and how hospitals are using this information of excess days to implement change. Moving to a vote, the Standing Committee passed the measure on usability.

The Standing Committee then voted to recommend the measure for continued endorsement. The Standing Committee observed that there are several related measures to this metric, but it did not consider these measures to be competing.

#3612 Risk-Standardized Acute Cardiovascular-Related Hospital Admission Rates for Patients With Heart Failure Under the Merit-Based Incentive Payment System – (Yale CORE)

Measure Steward/Developer Representatives at the Meeting

Kasia Lipska

Elizabeth Drye

Doris Peter

Lisa Marie Gomez

Jim Poyer (CMS)

Standing Committee Votes

- **Evidence**: Pass-14; No Pass-0
- **Performance Gap**: H-1; M-13; L-0; I-0
- **Reliability**: Yes-13; No-1
    - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.

  - The NQF Scientific Methods Panel's rating for Reliability: Moderate (H-0; M-5; L-3; I-0)
  - The Standing Committee accepted the NQF Scientific Methods Panel's rating.

- **Validity**: Yes-14; No-1
  - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.
  - The NQF Scientific Methods Panel's rating for Validity: Moderate (H-0; M-6; L-2; I-0)
  - The Standing Committee accepted the NQF Scientific Methods Panel's rating.

- **Feasibility**: H-5; M-9; L-0; I-0
- **Use**: Pass-14; No Pass-0
- **Usability**: H-0; M-13; L-1; I-0

Standing Committee Recommendation for Endorsement: Yes-13; No-1 (Pass – 13/14)

The Standing Committee recommended this outcome measure for endorsement.

The developer presented an overview of the measure, noting that this new measure assesses the risk-standardized rate of acute, unplanned, cardiovascular-related hospital admissions among Medicare Fee-for-Service (FFS) patients ages 65 years and older with HF or cardiomyopathy. The developer posits that there is strong evidence that ambulatory care clinicians can influence admission rates by providing high quality of care.

The Standing Committee reviewed and discussed the evidence, noting that the developer outlined a logic model depicting rates of admissions for patients with HF that can be decreased through care coordination and continuity of care from outpatient providers. The developer also cited evidence suggesting that outpatient clinicians can improve HF patients' risk of hospitalizations in a variety of ways, including, but not limited to, accessible primary care, coordination across providers and care settings, early attention to changes in clinical status, adoption of guideline-directed medical therapy, careful prescribing in patients with comorbidities, patient education, and support for self-management. Considering this information, the Standing Committee passed the measure on the evidence criterion.

In reviewing performance gap for this measure, the Standing Committee recognized that across all tax identification numbers (TINs), the cardiovascular-related risk-standardized acute admission rate (RSAAR) measure scores ranged from 9.6 to 62.4 per 100 person-years, with a median of 24.8 and an interquartile range (IQR) of 24.0 to 25.9. The mean RSAAR and standard deviation were 25.1 ± 2.4 admissions per 100 person-years. The Standing Committee also acknowledged that the distributions of RSCARs were generally similar with respect to the proportion of Medicare-Medicaid dual-eligible beneficiaries across TINs.

Moving to scientific acceptability, the Standing Committee recognized that the SMP reviewed and passed this measure on both reliability and validity. For reliability, the Standing Committee reviewed the testing data, in which the developer noted that a minimum reliability of 0.4 was achieved for TINs with at least 21 HF patients. At this threshold, reliability scores for TINs ranged from 0.40 to nearly 1.0, with a median value of 0.600 (IQR of 0.481-0.778). The SMP members agreed that the approach is appropriate; however, they raised several concerns, including clarity on the unit of analysis (clinician versus clinician group). One Standing Committee member raised concern with the low reliability results at a patient volume of 21 HF patients. The developer responded to the SMP's concerns, noting that under the Merit-Based Incentive Payment System (MIPS), clinicians annually select whether to report as individuals, as part of a group, or as both. The group includes both solo clinicians (i.e., clinicians opting not to report with other clinicians under MIPS) and groups of clinicians who have chosen to report their quality under

a common TIN. Therefore, testing results include both individual clinicians and clinician groups, which is consistent with how the MIPS program evaluates quality. Regarding the reliability results, the 21-minimum case volume was established to reach the reliability threshold of 0.4, which is acceptable for CMS; the MIPS program will set the minimum case volume during rulemaking. Considering this information, the Standing Committee agreed to uphold the SMP's rating on reliability.

For validity, the Standing Committee reviewed the validity testing for the measure. The developer conducted face validity of the measure score, which is the minimum acceptable testing for new measures. Of the 17 TEP members who were active through the end of the project, 12 of them responded. The majority of the respondents (10/12 or 83 percent) moderately or somewhat agreed that the MIPS HF measure can be used to distinguish *good* from *poor* quality of care. Of the 13 clinician Committee members who responded to the survey, 11 of them (85 percent) strongly, moderately, or somewhat agreed that the MIPS HF measure can be used to distinguish good from poor quality of care. The developer also adjusted for 30 risk variables, including the AHRQ SES Index. The r-squared value for the model with demographic and clinical risk factors was 0.073. The r-squared value after adding the AHRQ SES Index to the model was unchanged (0.073). The Standing Committee agreed that the concerns and discussions related the r-squared value for NQF #2880 apply to NQF #3612 as well. One Standing Committee member also mentioned that the more homogeneous the patient population is, such as HF patients, the less variation is seen within the model. Considering this information, the Standing Committee proceeded to vote to uphold the SMP's rating for validity.

The Standing Committee did not raise any concerns with respect to feasibility and passed the measure on this criterion. The Standing Committee also recognized that the measure is not currently publicly reported or used in an accountability application. However, CMS may propose this measure for use under MIPS. The Standing Committee did not raise any major concerns and passed the measure on use and usability.

Lastly, the Standing Committee voted to recommend the measure for continued endorsement. The Standing Committee observed that there is a related measure to this metric, but it did not consider this measure to be competing.

Public Comment

No public or NQF member comments were provided during the measure evaluation meeting.

Next Steps

Ms. Oroma Igwe, NQF manager, reviewed the next steps for the project, noting that NQF staff will incorporate the Standing Committee's spring 2021 measure evaluation discussion and voting results into the Spring 2021 Draft Technical Report. NQF will post the draft technical report on August 19, 2021, for a 30-calendar day public commenting period. The continuous public commenting period with member support will close on September 17, 2021. NQF will reconvene the Standing Committee for the post-comment web meeting on October 15, 2021, if it receives comments prior to the end of the draft report commenting period that necessitate discussion from the Standing Committee. Ms. Igwe also informed the Standing Committee that the Consensus Standards Approval Committee (CSAC) will consider the Standing Committee's endorsement recommendations during its meetings on November 30–December 1, 2021. Following the CSAC meeting, the 30-day Appeals period will be held from December 7–January 5, 2022.